The Annals of Statistics
1996, Vol. 24, No. 3, 1400–1422

THE BAHADUR–KIEFER REPRESENTATION FOR U-QUANTILES[1]

BY MIGUEL A. ARCONES

University of Texas

We consider the distributional and the almost sure pointwise Bahadur–Kiefer representation for U-quantiles. We show that the order of this representation depends on the order of the local variance of the empirical process of U-statistic structure at the U-quantile. Our results indicate that U-quantiles can be smoother than quantiles. U-quantiles can either be as unsmooth as quantiles or can behave as differentiable statistical functionals.

1. Introduction. First, let us recall the Bahadur–Kiefer representation for quantiles. Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. r.v.'s, let F_n be the empirical distribution function, let F be the cumulative distribution function of X_i and let $0 < p < 1$. Define $\xi_n := \inf\{t: F_n(t) \ge p\}$ and

$$R_n := \xi_n - \xi_0 + \left(F'(\xi_0)\right)^{-1}\left(F_n(\xi_0) - F(\xi_0)\right), \tag{1.1}$$

where $F(\xi_0) = p$. Kiefer (1967) showed that if F is second differentiable at ξ_0 and $F'(\xi_0) > 0$, then

$$n^{3/4}R_n \to_d p^{1/4}(1-p)^{1/4}|g_1|^{1/2}g_2, \tag{1.2}$$

where g_1 and g_2 are two independent standard normal r.v.'s. He also proved that

$$\limsup_{n\to\infty} \pm (n/2\log\log n)^{3/4}R_n = 2^{1/2}3^{-3/4}p^{1/4}(1-p)^{1/4} \quad \text{a.s.} \tag{1.3}$$

The purpose of this paper is to present similar results for U-quantiles. Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. r.v.'s with values in a measurable space $(S, \mathscr{S})$. Let h: $S^m \to \mathbb{R}$ be a measurable symmetric function. Let $H(t) = \Pr\{h(X_1, \dots, X_m) \le t\}$. The empirical distribution of U-statistic structure is defined by

$$H_n(t) := \frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_m)\in I_m^n} I_{h(X_{i_1},\dots,X_{i_m}) \le t}, \tag{1.4}$$

Received December 1993; revised July 1995.

[1]Research supported in part by NSF Grant DMS-93-02583 and carried out at the University of Utah.

AMS 1991 *subject classifications*. Primary 62E20; secondary 60F05, 60F15.

Key words and phrases. Quantiles, Bahadur–Kiefer representations, U-statistics, empirical processes.

where $I_m^n = \{(i_1, \dots, i_m): 1 \le i_j \le n \text{ and } i_k \ne i_j \text{ for } k \ne j\}$. Let $0 < p < 1$. Suppose that $H(\xi_0) = p$. The U-quantile is defined by

$$\xi_n := \inf\{t: H_n(t) \ge p\}. \tag{1.5}$$

Several common estimators are U-quantiles. For example, one very often used alternative to the median as a center of symmetry is the Hodges–Lehmann estimator: the median of $2^{-1}(X_i + X_j)$, $1 \le i < j \le n$ [see Hodges and Lehmann (1963)]. This is the U-quantile (with respect to $p = 1/2$) of the kernel $h(x_1, x_2) = 2^{-1}(x_1 + x_2)$. We refer to Lehmann (1975) for different extensions of this estimator and applications to nonparametric statistics. Another interesting example is the U-quantile of the kernel $h(x_1, x_2) = |x_1 - x_2|$ with respect to $p = 1/2$. This U-quantile is a measure of the spread of the distribution. It was introduced by Bickel and Lehmann (1979). Choudhury and Serfling (1988) introduced an U-quantile which estimates the regression slope. Consider the linear regression model: $Y_i = \alpha + \beta X_i + \delta_i$; α and β are constants and δ_i is an r.v. independent of X_i. The U-quantile of the kernel $h((x_1, y_1), (x_2, y_2)) = (y_2 - y_1)/(x_2 - x_1)$, with respect to $p = 1/2$, is a natural estimator of the parameter β. This estimator is the median of the values $(Y_j - Y_i)/(X_j - X_i)$, $1 \le i < j \le n$. Some references in the study of U-quantiles are Serfling (1984), Janssen, Serfling and Veraverbeke (1984), Helmers, Janssen and Serfling (1988) and Choudhury and Serfling (1988).

Here, we will study the distributional and the a.s. behavior of

$$R_n := \xi_n - \xi_0 + \left(H'(\xi_0)\right)^{-1}\left(H_n(\xi_0) - H(\xi_0)\right), \tag{1.6}$$

using empirical process techniques. Finding the asymptotic behavior of (1.6), we grasp a very good insight into the effect of the influence curve in the asymptotics of U-quantiles. The Bahadur–Kiefer representation of a statistical functional measures how close is, asymptotically, the linear expansion of a statistical functional to the statistical functional itself. It is a way to measure the differentiability of the statistical functional. We refer to Serfling (1980), Chapter 6, and Dudley (1992, 1994) for other ways to measure differentiability. One interesting application of Bahadur–Kiefer representations is to obtain sequential fixed-width confidence intervals for a parameter [see Chow and Robbins (1965) and Geertsema (1970)].

The leading idea to deal with (1.6) is to do a Hoeffding decomposition, to show that the terms of order 2 and larger vanish and to find the order from the first term of this decomposition. Next, we describe the Hoeffding decomposition. Given a measurable function on S^m, the U-statistic with kernel h is defined by

$$U_n(h) := \frac{(n-m)!}{n!} \sum_{(i_1, \dots, i_m) \in I_m^n} h(X_{i_1}, \dots, X_{i_m}). \tag{1.7}$$

We will abbreviate $Eh := E[h(X_1, \dots, X_m)]$, $P_n f = n^{-1}\sum_{j=1}^n f(X_j)$ and $Pf = E[f(X)]$, where X is a copy of X_1. We define

$$\pi_{k,m} f(x_1, \dots, x_k) = \left(\delta_{x_1} - P\right) \cdots \left(\delta_{x_k} - P\right) P^{m-k} f, \tag{1.8}$$

where $Q_1, \dots, Q_m f = \int \cdots \int f(x_1, \dots, x_m)\, dQ_1(x) \cdots dQ_m(x_m)$. Then, the Hoeffding decomposition of the U-statistic $U_n(f)$ can be written as

$$U_n(f) = \sum_{k=0}^{m} \binom{m}{k} U_n(\pi_{k,m} f). \tag{1.9}$$

In particular, this expansion applies to the term $(H'(\xi_0))^{-1}(H_n(\xi_0) - H(\xi_0))$ in (1.6), allowing us to see how close $\xi_n - \xi_0$ is to a true linear term. Here, we will see that the order of

$$E\Big[\big|g(X,t) - g(X,\xi_0)\big|^2\Big] \quad \text{as } t \to \xi_0 \tag{1.10}$$

determines the order of R_n, where $g(x,t) = \Pr\{h(x, X_2, \dots, X_m) \le t\}$. In particular, we will see that if $E[|g(X,t) - g(X,\xi_0)|^2] = O(|t - \xi_0|^v)$ as $t \to \xi_0$, for some $v > 0$, then

$$n^{(v+2)/4} R_n = O_{\mathrm{P}}(1) \tag{1.11}$$

and

$$(n/\log\log n)^{(v+2)/4} R_n = O(1) \quad \text{a.s.} \tag{1.12}$$

Observe that by the Cauchy–Schwarz inequality

$$\begin{aligned} |H(t) - H(\xi_0)|^2 &= \big|E\big[I_{h(X_1,\dots,X_m)\le t} - I_{h(X_1,\dots,X_m)\le \xi_0}\big]\big|^2 \\ &\le E_1\Big[\big|E_{2,\dots,m}\big[I_{h(X_1,\dots,X_m)\le t} - I_{h(X_1,\dots,X_m)\le \xi_0}\big]\big|^2\Big] \\ &= E\Big[\big|g(X,t) - g(X,\xi_0)\big|^2\Big] \\ &\le E\Big[\big|I_{h(X_1,\dots,X_m)\le t} - I_{h(X_1,\dots,X_m)\le \xi_0}\big|^2\Big] = |H(t) - H(\xi_0)|, \end{aligned} \tag{1.13}$$

where by $E_{i_1,\dots,i_k}$ we mean integration with respect to the coordinates $i_1, \dots, i_k$. So, if H is differentiable at ξ_0, $H'(\xi_0) > 0$ and $E[|g(X,t) - g(X,\xi_0)|^2] = O(|t - \xi_0|^v)$ as $t \to \xi_0$, for some $v > 0$, then $1 \le v \le 2$. Finding the exact order of (1.10) may be difficult or impossible, but, by (1.13), (1.11), and (1.12) always hold with $v = 1$. For a smooth statistical functional, the term R_n is $O_{\mathrm{P}}(n^{-1})$ and $O(n^{-1}(\log\log n))$ a.s. (case $v = 2$). These are the orders of all the examples mentioned above. These estimators enjoy a much better differentiability than the median.

We must mention the previous work in this problem. Choudhury and Serfling (1988) showed that, under some mild conditions,

$$n^{3/4}(\log n)^{-3/4}(H_n(\xi_0) - p + H'(\xi_0)(\xi_n - \xi_0)) = O(1) \quad \text{a.s.} \tag{1.14}$$

[see also Lemma 4.2 in Geertsema (1970)]. This result was used in Gijbels, Janssen and Veraverbeke (1988) to find weak and strong representations for trimmed U-statistics. We also must mention the work by Shi (1995) in the uniform Bahadur–Kiefer representation for the U-quantiles of $h(x_1, \dots, x_m) = \max(x_1, \dots, x_m)$. Other papers related to the present one are the ones by Carroll (1978), Jurečková (1980), Jurečková and Sen (1987), Deheuvels and Mason (1992) and Arcones (1994a) in the Bahadur–Kiefer representation of *M*-estimators.

Our main tools are certain limit theorems which hold uniformly over VC subgraph classes of functions. Given a set S and a collection of subsets $\mathscr{C}$, for $A \subset S$, let $\Delta^{\mathscr{C}}(A) = \text{card}\{A \cap C: C \in \mathscr{C}\}$, let $m^{\mathscr{C}}(n) = \max\{\Delta^{\mathscr{C}}(A): \text{card}(A) = n\}$ and let $s(\mathscr{C}) = \inf\{n: m^{\mathscr{C}}(n) < 2^n\}$; $\mathscr{C}$ is said to be a VC class of sets if $s(\mathscr{C}) < \infty$. General properties of VC classes of sets can be found in Chapters 9 and 11 in Dudley (1984). Given a function $f: S \to \mathbb{R}$, the subgraph of f is the set $\{(x, t) \in S \times \mathbb{R}: 0 \le t \le f(x) \text{ or } f(x) \le t \le 0\}$. A class of functions $\mathscr{F}$ is a VC subgraph class if the collection of subgraphs of $\mathscr{F}$ is a VC class. The interest of these classes of functions lies in their good properties with respect to covering numbers. Given a pseudometric space (T, d) the ε-covering number $N(\varepsilon, T, d)$ is defined by

$$N(\varepsilon, T, d) = \min\{m: \text{there exists a covering of } T \text{ by } m \text{ balls of radius} \le \varepsilon\}. \tag{1.15}$$

Given a positive measure μ on $(S, \mathscr{S})$, we define $N_2(\varepsilon, \mathscr{F}, \mu) = N(\varepsilon, \mathscr{F}, \|\cdot\|_{L_2(\mu)})$. If $\mathscr{F}$ is a VC subgraph class [Pollard (1984), Proposition 2.25], there are finite constants A and v such that, for each probability measure μ with $\mu F^2 < \infty$,

$$N_2(\varepsilon, \mathscr{F}, \mu) \le A\left(\left(\mu F^2\right)^{1/2}/\varepsilon\right)^v, \tag{1.16}$$

where $F(x) = \sup_{f \in \mathscr{F}}|f(x)|$ and A and v can be chosen depending only on $s(\mathscr{F})$, that is, uniformly over all the classes of functions with the same number $s(\mathscr{F})$. By the maximal inequality for sub-Gaussian processes [see Theorem 2.3.1 in Marcus and Pisier (1981); see also Theorem 1 in Dudley (1967)], there is a constant c depending only on A and v such that for any class of functions satisfying (1.16),

$$n^{-1}E\left[\sup_{f\in\mathscr{F}}\left|\sum_{i=1}^{n} \varepsilon_i f(X_i)\right|^2\right] \le cE\left[F^2(X)\right], \tag{1.17}$$

where $\{\varepsilon_i\}_{i=1}^{\infty}$ is a Rademacher sequence independent of the sequence $\{X_i\}_{i=1}^{\infty}$.

One of the main ingredients to study the distributional Bahadur–Kiefer representation of *U*-quantiles will be the weak convergence of a sequence of stochastic processes. By weak convergence, we mean weak convergence of random elements with values in $l_\infty(\mathscr{F})$ as in Hoffmann–Jørgensen (1984). $l_\infty(\mathscr{F})$ is the Banach space formed by all the uniformly bounded functions on $\mathscr{F}$ with the norm $\|x\|_{\mathscr{F}} = \sup_{f \in \mathscr{F}}|x(f)|$. Let $\{Z_n(f): f \in \mathscr{F}\}$, $n \ge 1$, be a sequence of stochastic processes, and let $\{Z(f): f \in \mathscr{F}\}$ be another stochastic process. The sequence of stochastic processes $\{Z_n(f): f \in \mathscr{F}\}$ is said to converge weakly to $\{Z(f): f \in \mathscr{F}\}$ in $l_\infty(\mathscr{F})$ if:

(i) $\sup_{f \in \mathscr{F}}|Z_n(f)| < \infty$ a.s. for each n large enough;
(ii) there exists a separable set S of $l_\infty(\mathscr{F})$ such that $\Pr^*\{Z \in S\} = 1$;
(iii) $E^*[H(Z_n)] \to E[H(Z)]$ for each bounded, continuous function H in $l_\infty(\mathscr{F})$.

It is well known [see, e.g., Andersen and Dobrić (1987)] that this type of convergence is equivalent to the convergence of the finite-dimensional distributions plus a finite-dimensional approximation, that is, $\{Z_n(f): f \in \mathscr{F}\}$, $n \geq 1$, converges weakly to $\{Z(f): f \in \mathscr{F}\}$ if and only if the finite-dimensional distributions of $\{Z_n(f): f \in \mathscr{F}\}$ converge to those of $\{Z(f): f \in \mathscr{F}\}$; and for each $\eta > 0$, there exists a map $\pi: \mathscr{F} \to \mathscr{F}$ such that $\#\{\pi f: f \in \mathscr{F}\}$ is finite and

$$\limsup_{n\to\infty} \Pr{}^*\Big\{\sup_{f\in\mathscr{F}} |Z_n(f) - Z_n(\pi f)| \geq \eta\Big\} \leq \eta. \tag{1.18}$$

We also will use the fact that if $\{Z(f): f \in \mathscr{F}\}$ is a stochastic process such that there exists a separable set S of $l_\infty(\mathscr{F})$ with $\Pr^*\{Z \in S\} = 1$, then $(\mathscr{F}, \rho_1)$ is totally bounded and $\Pr^*\{Z \in C_u(\mathscr{F}, \rho_1)\} = 1$, where $C_u(\mathscr{F}, \rho_1)$ is the set of all uniformly bounded and ρ_1-uniformly continuous functions in $\mathscr{F}$ and $\rho_1(f_1, f_2) := E[\min(|Z(f_1) - Z(f_2)|, 1)]$ [see Arcones (1995)]. In particular, if $\{Z_n(f): f \in \mathscr{F}\}$ converge weakly to $\{Z(f): f \in \mathscr{F}\}$, then

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \Pr{}^*\Bigg\{\sup_{\substack{\rho_1(f_1,f_2)\leq\delta\\ f_1,f_2\in\mathscr{F}}} |Z_n(f_1) - Z_n(f_2)| \geq \eta\Bigg\} = 0 \tag{1.19}$$

for each $\eta > 0$.

To study the almost sure Bahadur–Kiefer representation of U-quantiles instead of using weak convergence, we use a property similar to the compact law of the iterated logarithm: $\{Z_n(f): f \in \mathscr{F}\}$ is a sequence of stochastic processes such that there is a subset K of $l_\infty(\mathscr{F})$ satisfying that, with probability 1, $\{Z_n(f): f \in \mathscr{F}\}$ is relatively compact in $l_\infty(\mathscr{F})$ and its limit set is K. Given a sequence of stochastic processes $\{Z_n(f): f \in \mathscr{F}\}$ and a subset K of $l_\infty(\mathscr{F})$, we have that the following are equivalent:

(a) With probability 1, $\{Z_n(f): f \in \mathscr{F}\}$ is relatively compact in $l_\infty(\mathscr{F})$ and its limit set is K.

(b) For each $f_1, \dots, f_m \in \mathscr{F}$, with probability 1, $\{(Z_n(f_1), \dots, Z_n(f_m))\}$, $n \geq 1$, is relatively compact in $\mathbb{R}^m$ and its limit set is $\{(x(f_1), \dots, x(f_m)): x \in K\}$; and for each $\eta > 0$, there exists a map $\pi: \mathscr{F} \to \mathscr{F}$ such that $\#\{\pi f: f \in \mathscr{F}\}$ is finite and

$$\limsup_{n\to\infty} \sup_{f\in\mathscr{F}} |Z_n(f) - Z_n(\pi f)| \leq \eta \quad \text{a.s.} \tag{1.20}$$

If either (a) or (b) holds holds, K is a compact set of $l_\infty(\mathscr{F})$. If K is a compact set of $l_\infty(\mathscr{F})$, then $(\mathscr{F}, \rho_2)$ is totally bounded and $K \subset C_u(\mathscr{F}, \rho_2)$, where $\rho_2(f_1, f_2) := \sup_{x\in K} |x(f_1) - x(f_2)|$. In particular,

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \sup_{\substack{\rho_2(f_1,f_2)\leq\delta\\ f_1,f_2\in\mathscr{F}}} |Z_n(f_1) - Z_n(f_2)| = 0 \quad \text{a.s.} \tag{1.21}$$

We refer for all last facts on the compact law of the iterated logarithm to Arcones and Giné (1995). Usually K is the unit ball of the reproducing kernel Hilbert space (r.k.h.s.) of a covariance function on $\mathscr{F}$. A function $R: \mathscr{F} \times \mathscr{F} \to \mathbb{R}$

is a covariance function on $\mathscr{F}$, if

$$\sum_{j=1}^{m}\sum_{k=1}^{m} a_j a_k R(f_j, f_k) \geq 0$$

for each $a_1, \dots, a_m \in \mathbb{R}$ and each $f_1, \dots, f_m \in \mathscr{F}$. Then there is a mean-zero Gaussian process $\{W(f)\colon F \in \mathscr{F}\}$ such that $E[W(f_1)W(f_2)] = R(f_1, f_2)$ for each $f_1, f_2 \in \mathscr{F}$. Let $\mathscr{L}$ be the linear subspace of L_2, generated by $\{W(f)\colon f \in \mathscr{F}\}$. The r.k.h.s. of the covariance function $R(\cdot, \cdot)$ is the following class of functions on $\mathscr{F}$:

$$\big\{(\alpha(f))_{f\in\mathscr{F}}\colon \text{there exists } \xi \in \mathscr{L} \text{ such that } \alpha(f) = E[W(f)\xi] \text{ for each } f \in \mathscr{F}\big\}.$$

This space is endowed with the inner product

$$\langle \alpha_1, \alpha_2 \rangle := E[\xi_1 \xi_2]$$

if $\alpha_i(f) = E[W(f)\xi_i]$ for each $f \in \mathscr{F}$ and each $i = 1, 2$. The unit ball of this r.k.h.s. is

$$K := \big\{(E[W(f)\xi])_{f\in\mathscr{F}}\colon \xi \in \mathscr{L} \text{ and } E[\xi^2] \leq 1\big\}.$$

Here, $\rho_2(f_1, f_2) = \sup_{x \in K}|x(f_1) - x(f_2)| = \|W(f_1) - W(f_2)\|_2$. A reference in r.k.h.s.'s is Aronszajn (1950).

2. The distributional Bahadur–Kiefer representation for *U*-quantiles. Here, we consider the distributional order of the Bahadur–Kiefer representation for *U*-quantiles. First, we give an upper bound to R_n.

THEOREM 1. *With the above notation, suppose that*:

(i) *there is a real number* ξ_0 *such that* $H(\xi_0) = p$, *H is continuous in a neighborhood of* ξ_0, *H is differentiable at* ξ_0, $H'(\xi_0) > 0$ *and*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + O\big((t - \xi_0)^2\big)$$

as $t \to \xi_0$;

(ii) *there is a sequence of real numbers* $\{a_n\}_{n=1}^{\infty}$ *such that*

$$a_n^2 E\Big[\big|g(X, \xi_0 + tn^{-1/2}) - g(X, \xi_0)\big|^2\Big] = O(1)$$

for each $t \in \mathbb{R}$.

Then

$$a_n n^{1/2}(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)) = O_{\mathrm{P}}(1). \tag{2.1}$$

PROOF. c will design a finite constant which may vary from line to line. By (1.13) and hypotheses (i) and (ii),

$$a_n^2 n^{-1} = O(1). \tag{2.2}$$

Since, for each $\varepsilon > 0$, $H_n(\xi_0 \pm \varepsilon) \to H(\xi_0 \pm \varepsilon)$ a.s. and $H(\xi_0 - \varepsilon) < p < H(\xi_0 + \varepsilon)$, we have that

$$\xi_n \to \xi_0 \quad \text{a.s.} \tag{2.3}$$

Since the class $\{I_{h(X_1,\dots,X_m)\le t}: t \in \mathbb{R}\}$ is a VC subgraph class of functions, by Theorem 4.10 in Arcones and Giné (1993), $\{n^{1/2}(H_n(t) - H(t)): t \in \mathbb{R}\}$ converges weakly to a Gaussian process. From this, (2.3) and the fact that $E[|g(X,t) - g(X,\varepsilon_0)|^2] \to 0$, as $t \to \xi_0$, it follows that

$$n^{1/2}(H_n(\xi_n) - H_n(\xi_0) - H(\xi_n) + H(\xi_0)) \to_{\mathrm{P}} 0. \tag{2.4}$$

By hypothesis (i), there is a $\delta > 0$ such that $H(t)$ is continuous and increasing in $[\xi_0 - \delta, \xi_0 + \delta]$. Hence, for $\{i_1,\dots,i_m\} \cap \{j_1,\dots,j_m\} = \varnothing$ and $i_1 < \cdots < i_m$ and $j_1 < \cdots < j_m$,

$$\Pr\{h(X_{i_1},\dots,X_{i_m}) = h(X_{j_1},\dots,X_{j_m}) \in [\xi_0 - \delta, \xi_0 + \delta]\} = 0.$$

This implies that, for all $|s - \xi_0| < \delta$,

$$|H_n(s) - H_n(s-)| \le \binom{n}{m}^{-1}\left|\binom{n}{m} - \binom{n-m}{m}\right| \le cn^{-1} \quad \text{a.s.}$$

Therefore, eventually

$$|H_n(\xi_n) - p| \le cn^{-1} \quad \text{a.s.} \tag{2.5}$$

By (2.4) and (2.5),

$$n^{1/2}(H_n(\xi_0) - H(\xi_0) + H(\xi_n) - H(\xi_0)) \to_{\mathrm{P}} 0. \tag{2.6}$$

By hypothesis (i), there exists a positive constant η such that if $|t - \xi_0| \le \eta$, then $|H(t) - H(\xi_0)| \ge 2^{-1}H'(\xi_0)|t - \xi_0|$. So, if $|\xi_n - \xi_0| \le \eta$, then

$$\begin{aligned} 2^{-1}H'(\xi_0)n^{1/2}|\xi_n - \xi_0| &\le n^{1/2}|H(\xi_n) - H(\xi_0)| \\ &\le n^{1/2}|H_n(\xi_0) - H(\xi_0) + H(\xi_n) - H(\xi_0)| \\ &\quad + n^{1/2}|H_n(\xi_0) - H(\xi_0)| = O_{\mathrm{P}}(1). \end{aligned}$$

From this and (2.3), $n^{1/2}|\xi_n - \xi_0| = O_{\mathrm{Pr}}(1)$. The last estimation, (2.6) and hypothesis (i) imply that

$$n^{1/2}(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)) \to_{\mathrm{P}} 0.$$

Next, we show that, for each $M < \infty$,

$$\begin{aligned} \sup_{|t|\le M} a_n n^{1/2}\big|H_n(\xi_0 + tn^{-1/2}) - H(\xi_0 + tn^{-1/2}) \\ - H_n(\xi_0) + H(\xi_0)\big| = O_{\mathrm{P}}(1). \end{aligned} \tag{2.7}$$

By the Hoeffding decomposition, it suffices to show that

$$\sup_{|t|\le M} a_n n^{1/2}\big|(P_n - P)\big(g(\cdot, \xi_0 + tn^{-1/2}) - g(\cdot, \xi_0)\big)\big| = O_{\mathrm{P}}(1) \tag{2.8}$$

and

$$\sup_{|t|\le M} a_n n^{1/2}\big|U_n \pi_{k,m}(I_{h\le \xi_0 + tn^{-1/2}} - I_{h\le \xi_0})\big| \to_{\mathrm{P}} 0 \tag{2.9}$$

for $k \le 2 \le m$. Since the class $\{g(x,t)\colon t \in \mathbb{R}\}$ is increasing in $\mathbb{R}$, it is a VC subgraph class. Hence, by (1.17),

$$E\left[a_n^2 n \sup_{|t|\le M} \left|(P_n - P)\big(g(\cdot, \xi_0 + tn^{-1/2}) - g(\cdot, \xi_0)\big)\right|^2\right]$$
$$\le c a_n^2 E\left[\sup_{|t|\le M} \left|g(X, \xi_0 + tn^{-1/2}) - g(X, \xi_0)\right|^2\right]$$
$$\le c a_n^2 E\left[\left|g(X, \xi_0 + Mn^{-1/2}) - g(X, \xi_0)\right|^2\right]$$
$$+ a_n^2 E\left[\left|g(X, \xi_0 - Mn^{-1/2}) - g(X, \xi_0)\right|^2\right] = O(1).$$

[Observe that the classes $\mathcal{F}_n := \{g(\cdot, \xi_0 + tn^{-1/2}) - g(\cdot, \xi_0)\colon |t| \le M\}$ are all VC subgraph classes and $s(\mathcal{F}_n) \le s(\mathcal{F}_1)$ for each n. So, we may choose a finite constant c in (1.17) uniformly on n.] Therefore, (2.8) follows. (2.9) follows from (2.2) and Corollary 5.7 in Arcones and Giné (1993). Hence, (2.7) holds. Composing the process in (2.7) with $n^{1/2}(\xi_n - \xi_0)$, we get that

$$a_n n^{1/2}\big(H_n(\xi_n) - H(\xi_n) - H_n(\xi_0) + H(\xi_0)\big) = O_{\mathrm{P}}(1).$$

By this, (2.2) and (2.5),

$$a_n n^{1/2}\big(H(\xi_0) - H(\xi_n) - H_n(\xi_0) + H(\xi_0)\big) = O_{\mathrm{P}}(1).$$

So, the result follows. □

From a previous theorem, it follows that if condition (i) holds and there exists a $1 \le v \le 2$ such that $E[|g(X,t) - g(X,\xi_0)|^2] = O(|t - \xi_0|^v)$ as $t \to \xi_0$, then

$$n^{(v+2)/4} R_n = O_{\mathrm{P}}(1).$$

Next, we will find the exact order of this representation under some extra conditions. The exact order of this representation is determined by the order of

$$E\left[\left|g(X, \xi_0 + tn^{-1/2}) - g(X, \xi_0)\right|^2\right] \tag{2.10}$$

as $n \to \infty$. Finding the order of (2.10) could be difficult. By (1.13),

$$n^{1/2} E\left[\left|g(X, \xi_0 + tn^{-1/2}) - g(X, \xi_0)\right|^2\right]$$
$$\le n^{1/2}\left|H(\xi_0 + tn^{-1/2}) - H(\xi_0)\right| = O(1),$$

that is, condition (ii) holds with $a_n = n^{1/4}$. So, under the easy-to-verify condition (i) in Theorem 1, we have that

$$n^{3/4}\big(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)\big) = O_{\mathrm{P}}(1). \tag{2.11}$$

We will need the following CLT for triangular arrays indexed by VC classes. It follows from Theorem 2.6 in Alexander (1987).

THEOREM 2. *Let Θ be a subset of $\mathbb{R}^d$ and let θ_0 be a point in the interior of Θ. Let g: $S \times \Theta \to \mathbb{R}$ be a function such that $g(\cdot, \theta)$ is a measurable function for each $\theta \in \Theta$. Let $M < \infty$. Let $G_R(x) = \sup_{|\theta-\theta_0| \le R} |g(x,\theta) - g(x,\theta_0)|$, where $|\cdot|$ is the Euclidean distance $\mathbb{R}^d$. Suppose that*:

(i) *there is a $\delta_0 > 0$ such that $\{g(x,\theta) - g(x,\theta_0): |\theta - \theta_0| \le \delta_0\}$ is a VC subgraph class*;

(ii) *there are sequences $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ such that $a_n \to \infty$, $b_n \to 0$ and*

$$\lim_{n\to\infty} a_n^2 \operatorname{Var}\big(g(X, \theta_0 + tb_n) - g(X, \theta_0 + sb_n)\big)$$

exists for each $|s|, |t| \le M$;

(iii) $a_n^2 E[G_{Mb_n}^2(X)] = O(1)$;

(iv) $a_n^2 E[G_{Mb_n}^2(X) I_{G_{Mb_n}(X) \ge \tau a_n^{-1} n^{1/2}}] \to 0$ *for each* $\tau > 0$;

(v) $\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{|t-s|\le\delta, |s|,|t|\le M} a_n \|g(X,\theta_0 + tb_n) - g(X, \theta_0 + sb_n)\|_2 = 0$.

Then

$$\bigg\{a_n n^{-1/2} \sum_{i=1}^{n} \big(g(X_i, \theta_0 + tb_n) - g(X_i, \theta_0) - E\big[g(X, \theta_0 + tb_n)\big] + E\big[g(X, \theta_0)\big]\big): |t| \le M\bigg\}$$

converges weakly to the centered Gaussian process $\{Z(t): |t| \le M\}$ determined by $Z(0) = 0$ and $\|Z(t) - Z(s)\|_2 = \rho(t,s)$.

Next, we see that under some conditions the order $a_n = n^{1/4}$ is attained.

THEOREM 3. *Suppose that*:

(i) *there is a real number ξ_0 such that $H(\xi_0) = p$, H is continuous in a neighborhood of ξ_0, H is differentiable at ξ_0, $H'(\xi_0) > 0$ and there exists a finite constant b such that*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + b(t-\xi_0)^2 + o\big((t-\xi_0)^2\big)$$

as $t \to \xi_0$;

(ii) *there is a real number β such that*

$$\lim_{\varepsilon\to 0+} \varepsilon^{-1} E\Big[\big|g(X, \xi_0 + \varepsilon t) - g(X, \xi_0 + \varepsilon s)\big|^2\Big] = \beta^2 |t-s|$$

for each $t, s \in \mathbb{R}$.

Then

$$n^{3/4}\big(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)\big) \to_d m^{3/2} \beta \big(H'(\xi_0)\big)^{-1/2} \big(\operatorname{Var}(g(X, \xi_0))\big)^{1/4} |g_1|^{1/2} g_2,$$

where g_1 and g_2 are two independent standard normal r.v.'s.

PROOF. We claim that by Theorem 2,

$$\{n^{3/4}m(P_n - P)\big(g(\cdot, \xi_0 + tn^{-1/2}) - g(\cdot, \xi_0)\big): |t| \le M\big)\} \tag{2.12}$$

converges weakly to $\{Z(t): |t| \le M\}$, where Z is a mean-zero Gaussian process with covariance given by

$$E[Z(t)Z(s)] = 2^{-1}m^2\beta(|t| + |s| - |t - s|)$$

for each $s, t \in \mathbb{R}$. Observe that $g(x, t)$ is nondecreasing in t for each fixed x. So, by a standard approximation argument,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{\substack{|t-s|\le\delta \\ |s|, |t|\le M}} n^{1/2}E\Big[\big|g(X, \xi_0 + tn^{-1/2}) - g(X, \xi_0 + sn^{-1/2})\big|^2\Big] = 0$$

for each $M < \infty$. Observe also that $|g(x, t)| \le 1$.

By Corollary 5.7 in Arcones and Giné (1993),

$$\sup_{|t|\le M} n^{3/4}\big|U_n\pi_{k,m}\big(I_{h\le \xi_0 + tn^{-1/2}} - I_{h\le\xi_0}\big)\big| \to_{\mathrm{P}} 0 \quad \text{for } 2 \le k \le m. \tag{2.13}$$

From (2.12) and (2.13), the process

$$\begin{aligned}\{Z_n(t) := n^{3/4}\big(H_n(\xi_0 + tn^{-1/2}) - H(\xi_0 + tn^{-1/2})\\ -H_n(\xi_0) + H(\xi_0)\big): |t| \le M\}\end{aligned}$$

converges weakly to the process $\{Z(t): |t| \le M\}$. Let $T_M^* = \{t: |t| \le M\} \cup \{\infty\}$, let $Z_n(\infty) := n^{1/2}(\xi_n - \xi_0)$, let $Z(\infty)$ be a Gaussian r.v. with mean zero, covariance $m^2(H'(\xi_0))^{-2}\operatorname{Var}(g(X, \xi_0))$ and independent of the process $\{Z(t): |t| \le M\}$. By (2.11)

$$n^{1/2}(\xi_n - \xi_0) + n^{1/2}(H'(\xi_0))^{-1}(N_n(\xi_0) - H(\xi_0)) \to_{\mathrm{P}} 0. \tag{2.14}$$

So, from this and the central limit theorem for triangular arrays, the finite-dimensional distributions of $\{Z_n(t): t \in T_M^*\}$ converge to those of $\{Z(t): t \in T_M^*\}$. Condition (1.18) holds for $\{Z_n(t): t \in T_M^*\}$, because it holds for $\{Z_n(t): |t| \le M\}$. Therefore, $\{Z_n(t): t \in T_M^*\}$ converges weakly to $\{Z(t): t \in T_M^*\}$.

By composing $Z_n(t)$ with $Z_n(\infty)$, we get that

$$n^{3/4}(H_n(\xi_n) - H(\xi_n) - H_n(\xi_0) + H(\xi_0)) \to_{\mathrm{d}} Z(Z(\infty)).$$

We have that $Z(\infty)$ has the distribution of $m(H'(\xi_0))^{-1}(\operatorname{Var}(g(X, \xi_0)))^{1/2}g_1$, where g_1 is a standard normal r.v. For each $t \in \mathbb{R}$, the distribution of

$$Z\Big(m(H'(\xi_0))^{-1}\big(\operatorname{Var}(g(X, \xi_0))\big)^{1/2}t\Big)$$

is that of

$$m\beta m^{1/2}(H'(\xi_0))^{-1/2}\big(\operatorname{Var}(g(X, \xi_0))\big)^{1/4}|t|^{1/2}g_2,$$

where g_2 is a standard normal r.v. So, by conditioning on g_1, we obtain that the distribution of

$$Z\Big(m(H'(\xi_0))^{-1}\big(\operatorname{Var}(g(X, \xi_0))\big)^{1/2}g_1\Big)$$

is that of

$$m^{3/2}\beta(H'(\xi_0))^{-1/2}(\mathrm{Var}(g(X,\xi_0)))^{1/4}|g_1|^{1/2}g_2,$$

where g_1 and g_2 are independent standard normal r.v.'s. □

Theorem 3 applies to $h(x_1,\dots,x_m)=\max_{1\le i\le m}x_i$. Let $F(t)$ be the distribution function of X_i. Suppose that $(F(\xi_0))^m=p$, F is second differentiable at ξ_0 and $F'(\xi_0)>0$. Then

$$n^{3/4}(H_n(\xi_0)-H(\xi_0)+H'(\xi_0)(\xi_n-\xi_0))$$
$$\to_d mp^{(4m-3)/4m}(1-p^{1/m})^{1/4}|g_1|^{1/2}g_2,$$

where g_1 and g_2 are independent standard normal r.v.'s. Observe that $H(t)=(F(t))^m$ and

$$g(x,t)=\Pr\{\max(x,X_2,\dots,X_m)\le t\}=I_{x\le t}(F(t))^{m-1}.$$

So, $H'(\xi_0)=mp^{(m-1)/m}F'(\xi_0)$ and $\mathrm{Var}(g(X,\xi_0))=p^{(2m-1)/m}(1-p^{1/m})$. For $s<t$

$$\epsilon^{-1}E\left[|g(X,\xi_0+t\epsilon)-g(X,\xi_0+s\epsilon)|^2\right]$$
$$=\epsilon^{-1}\left((F(\xi_0+t\epsilon))^{m-1}-F(\xi_0+s\epsilon))^{m-1}\right)^2F(\xi_0+s\epsilon)$$
$$+\epsilon^{-1}(F(\xi_0+t\epsilon))^{2m-2}(F(\xi_0+t\epsilon)-F(\xi_0+\epsilon s)),$$

which converges to $p^{2(m-1)/m}(t-s)F'(\xi_0)$. We also have that

$$\epsilon^{-1/2}E\left[(g(X,\xi_0+t\epsilon)-g(X,\xi_0))g(X,\xi_0)\right]\to 0.$$

Next, we see how the order n in the Bahadur–Kiefer representation of U-quantiles can be attained.

THEOREM 4. *Suppose that*:

(i) *there is a real number* ξ_0 *such that* $H(\xi_0)=p$, H *is continuous in a neighborhood of* ξ_0, $H'(\xi_0)>0$ *and there exists a finite constant b such that*

$$H(t)=H(\xi_0)+H'(\xi_0)(t-\xi_0)+b(t-\xi_0)^2+o\left((t-\xi_0)^2\right)$$

as $t\to\xi_0$ *for some* $b\in\mathbb{R}$;

(ii) *there is a real number* β *such that*

$$\lim_{\varepsilon\to 0+}\varepsilon^{-2}\,\mathrm{Var}(g(X,\xi_0+\varepsilon t)-g(X,\xi_0+\varepsilon s))=\beta(t-s)^2$$

for each $t,s\in\mathbb{R}$;

(iii) $$\lim_{\varepsilon\to 0+}\varepsilon^{-2}E\left[|g(X,\xi_0+\varepsilon t)-g(X,\xi_0)|^2 I_{|g(X,\xi_0+\varepsilon t)-g(X,\xi_0)|\ge\tau}\right]=0$$

for each $t\in\mathbb{R}$ *and each* $\tau>0$;

(iv) *there is a real number* α *such that*

$$\lim_{\epsilon \to 0+} \epsilon^{-1}\operatorname{Cov}(g(X, \xi_0 + \epsilon t) - g(X, \xi_0), g(X, \xi_0)) = \alpha t$$

for each $t \in \mathbb{R}$.

(v) *There is a* $\delta > 0$ *such that for any* $|t - \xi_0| < \delta$ *and for any combinations* $i_1 < \cdots < i_m$ *and* $j_1 < \cdots < j_m$ *such that* $\{i_1, \dots, i_m\} \neq \{j_1, \dots, j_m\}$, *we have that*

$$\Pr\left\{h\left(X_{i_1}, \dots, X_{i_m}\right) = h\left(X_{j_1}, \dots, X_{j_m}\right) = t\right\} = 0.$$

Then

$$n(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0))$$

converges in distribution to $\lambda_1 g_1^2 + \lambda_2 g_2^2$, *where* g_1 *and* g_2 *are independent standard normal random variables*,

$$\lambda_1 = -2^{-1}(c_{1,2} + bc_{2,2}) - 2^{-1}\left(c_{1,1}c_{2,2} + 2bc_{1,2}c_{2,2} + b^2c_{2,2}^2\right)^{1/2}, \tag{2.15}$$

$$\lambda_2 = -2^{-1}(c_{1,2} + bc_{2,2}) + 2^{-1}\left(c_{1,1}c_{2,2} + 2bc_{1,2}c_{2,2} + b^2c_{2,2}^2\right)^{1/2}, \tag{2.16}$$

$$\begin{aligned} &c_{1,1} = m^2\beta^2, c_{2,2} = m^2\left(H'(\xi_0)\right)^{-2}\operatorname{Var}(g(X, \xi_0)) \quad \text{and} \\ &c_{1,2} = -m^2\left(H'(\xi_0)\right)^{-1}\alpha. \end{aligned} \tag{2.17}$$

PROOF. Observe that, by hypotheses (i) and (ii), $m \geq 2$. By the method in the proof of Theorem 3,

$$\left\{n\left(H_n\left(\xi_0 + tn^{-1/2}\right) - H\left(\xi_0 + tn^{-1/2}\right) - H_n(\xi_0) + H(\xi_0)\right): |t| \leq M\right\}$$

and $n^{1/2}(\xi_n - \xi_0)$ converge jointly to $\{tY_1 : |t| \leq M\}$ and Y_2, where Y_1 and Y_2 are jointly normal random variables with mean zero and

$$E\left[Y_1^2\right] = c_{1,1}, E\left[Y_2^2\right] = c_{2,2} \quad \text{and} \quad E[Y_1Y_2] = c_{1,2},$$

where $c_{1,1}$, $c_{2,2}$ and $c_{1,2}$ are as in (2.17). Hence

$$n\left(H_n(\xi_n) - H(\xi_n) - H_n(\xi_0) + H(\xi_0) + b(\xi_n - \xi_0)^2\right)$$

converges in distribution to Y_1, $Y_2 + bY_2^2$. We have that all $h(X_{i_1}, \dots, X_{i_m})$ in $[\xi_0 - t, \xi_0 + t]$ are different. So, $|H_n(\xi_n) - p| \leq \binom{n}{m}^{-1}$. Therefore,

$$n(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0))$$

converges in distribution to $-Y_1Y_2 - bY_2^2$. If $c_{1,1}c_{2,2} - c_{1,2}^2 = 0$, Y_1 and Y_2 are linearly dependent, and the distribution of $-Y_1Y_2 - bY_2^2$ is that of λg^2, where g is a standard normal random variable and

$$\lambda = E\left[-Y_1Y_2 - bY_2^2\right] = -(c_{1,2} + bc_{2,2}).$$

If $c_{1,1}c_{2,2} - c_{1,2}^2 = 0$, then

$$\lambda_1 g_1^2 + \lambda_2 g_2^2 = -(c_{1,2} + bc_{2,2})^+ g_1^2 - (c_{1,2} + bc_{2,2})^- g_2^2,$$

where $x^- = \max(-x, 0)$, $x^+ = \max(x, 0)$, λ_1 and λ_2 are as in (2.15) and (2.16) and g_1 and g_2 are two independent standard normal random variables. So, $\lambda_1 g_1^2 + \lambda_2 g_2^2$ has the same distribution as λg^2. Assume now that $c_{1,1}c_{2,2} - c_{1,2}^2 > 0$. Let $Z_1 = c_{2,2}^{-1/2}(c_{1,1}c_{2,2} - c_{1,2}^2)^{-1/2}(c_{2,2}Y_2 - c_{1,2}Y_2)$ and $Z_2 = c_{2,2}^{-1/2}Y_2$. We have that Z_1 and Z_2 are two independent standard normal random variables and $-Y_1Y_2 - bY_2^2 = -a_1Z_1Z_2 - a_2Z_2^2$, where $a_1 = (c_{1,1}c_{2,2} - c_{1,2}^2)^{1/2}$ and $a_2 = c_{1,2} + bc_{2,2}$. Let $g_1 = q_1^{-1}a_1Z_1 + q_1^{-1}(a_2 + (a_1^2 + a_2^2)^{1/2})Z_2$ and $g_2 = q_2^{-1}a_1Z_1 + q_2^{-1}(a_2 - (a_1^2 + a_2^2)^{1/2})Z_2$, where

$$
\begin{aligned}
q_1 &= 2^{1/2}\Big(a_1^2 + a_2^2 + a_2\big(a_1^2 + a_2^2\big)^{1/2}\Big)^{1/2} \quad \text{and} \\
q_2 &= 2^{1/2}\Big(a_1^2 + a_2^2 - a_2\big(a_1^2 + a_2^2\big)^{1/2}\Big)^{1/2}.
\end{aligned} \tag{2.18}
$$

Then, g_1 and g_2 are two independent standard normal random variables and

$$-a_1Z_1Z_2 - a_2Z_2^2 = \lambda_1 g_1^2 + \lambda_2 g_2^2,$$

where λ_1 and λ_2 are as in (2.15) and (2.16). So, the result follows. □

Theorem 4 applies to $h(x_1, x_2) = x_1 + x_2$. Suppose that $\Pr\{X_1 + X_2 \le \xi_0\} = p$, the distribution function F of X is second differentiable and its first and second derivatives are uniformly bounded and $\int_{-\infty}^{\infty} F'(\xi_0 - x)F'(x)\,dx > 0$. Then the thesis of Theorem 4 holds for $h(x_1, x_2) = x_1 + x_2$ with

$$
\begin{aligned}
H'(\xi_0) &= \int_{-\infty}^{\infty} F'(\xi_0 - x)F'(x)\,dx, \\
b &= 2^{-1}\int_{-\infty}^{\infty} F''(\xi_0 - x)F'(x)\,dx, \\
\beta^2 &= \int_{-\infty}^{\infty} (F'(\xi_0 - x))^2 F'(x)\,dx - \left(\int_{-\infty}^{\infty} F'(\xi_0 - x)F'(x)\,dx\right)^2, \\
\alpha &= \int_{-\infty}^{\infty} F'(\xi_0 - x)F(\xi_0 - x)F'(x)\,dx \\
&\quad - \int_{-\infty}^{\infty} F'(\xi_0 - x)F'(x)\,dx \int_{-\infty}^{\infty} F(\xi_0 - x)F'(x)\,dx
\end{aligned}
$$

and $g(x, t) = F(t - x)$.

Theorem 4 also applies to the kernel $h(x_1, x_2) = |x_1 - x_2|^r$, where $r > 0$. We omit the details. As mentioned in the Introduction, the U-quantile over this kernel, in the case $r = 1$ and $p = 1/2$, was considered by Bickel and Lehmann (1979). It is an estimator of the spread of the distribution.

Theorem 4 also applies to the kernel $h((x_1, y_1), (x_2, y_2)) = (y_2 - y_1)/(x_2 - x_1)$ (under some regularity conditions). Consider the linear regression model: $Y_i = \alpha + \beta X_i + \delta_i$; α and β are constants and δ_i is an r.v. independent of X_i. The U-quantile over the kernel h, with respect to $p = 1/2$, is an estimator of the regression slope β. Here

$$g((x_1, y_1), t) = \Pr\{(\beta - t)X_2 + \delta_2 \le y_1 - tx_1 - \alpha\}.$$

3. The a.s. Bahadur–Kiefer representation of *U*-quantiles. First, we present some results on the law of the iterated logarithm for processes. The following lemma is similar in spirit to Theorem 3.1 in Kuelbs (1976).

LEMMA 5. *Let $\{Z_n(t): t \in T\}$ be a stochastic process indexed by a parameter set T. Let ρ be a pseudometric on T. Let K be a compact subset of $C_u(T, \rho)$. Assume that the following conditions are satisfied*:

(i) *(T, ρ) is totally bounded*;

(ii) $\lim_{\delta \to 0} \limsup_{n\to\infty} \sup_{\rho(t_1, t_2) \le \delta} |Z_n(t_1) - Z_n(t_2)| = 0$ *a.s.*;

(iii) *for each $m \in \mathbb{N}$ and each $t_1, \dots, t_m \in T$, with probability* 1, *the sequence $\{(Z_n(t_1), \dots, Z_n(t_m))\}_{n=1}^{\infty}$ is relatively compact in $\mathbb{R}^m$ and its limit set is contained in $\{(x(t_1), \dots, x(t_m)): x \in K\}$.*

Then, with probability 1, *the sequence $\{Z_n(t): t \in T\}$ is relatively compact in $l_\infty(T)$ and its limit set is contained in K.*

PROOF. By the Arzelá–Ascoli theorem, with probability 1 $\{Z_n(t): t \in T\}$ is relatively compact in $l_\infty(T)$. Let $\{t_p\}_{p=1}^{\infty}$ be a countable dense subset of (T, ρ). Let A be a measurable set having probability 1, such that in A,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{\rho(t_1, t_2)\le\delta} \left| Z_n(t_1) - Z_n(t_2) \right| = 0$$

and for each $m \in \mathbb{N}$ the sequence $\{(Z_n(t_1), \dots, Z_n(t_m))\}_{n=1}^{\infty}$ is relatively compact in $\mathbb{R}^m$ and its limit set is contained in $\{(x(t_1), \dots, x(t_m)): x \in K\}$. Suppose that $x \in l_\infty(T)$ is a limit point of a sequence $\{Z_n(t): t \in T\}$ satisfying the previous two conditions. By the first condition, $x \in C_u(T, \rho)$. Let n_j be a subsequence such that $Z_{n_j} \to x$. By the second condition, for each $m \ge 1$, there is an $x^{(m)} \in K$ such that

$$\left(x(t_1), \dots, x(t_m)\right) = \left(x^{(m)}(t_1), \dots, x^{(m)}(t_m)\right).$$

Since K is compact, the sequence $\{x^{(m)}\}_{m=1}^{\infty}$ has a limit point $y \in K$. Since $x, y \in C_u(T, \rho)$ and $x(t_p) = y(t_p)$ for each $p \ge 1$, $x = y$. □

We also need the following law of the iterated logarithm.

THEOREM 6. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d. r.v.'s with values in a measurable space $(S, \mathscr{S})$. Let $g: S \times T \to \mathbb{R}$ be a function such that $g(\cdot, t): S \to \mathbb{R}$ is a measurable function for each $t \in T$. Let $R(\cdot, \cdot)$ be a covariance function on T. Let $\{b_n\}$ be a sequence of real numbers from the interval $(0, 1]$ and let $\{a_n\}$ be a sequence of positive real numbers. Suppose that*:

(i) *there is a scalar product defined for each $t \in T$ and each $0 \le u \le 1$, so that $ut \in T$*;

(ii) $\limsup_{n\to\infty} a_n^2 n^{-1} \operatorname{Var}(\sum_{j=1}^{p} \lambda_j g(X, b_n t_j)) \le \sum_{j,k=1}^{p} \lambda_j \lambda_k R(t_j, t_k)$ *for each $t_1, \dots, t_p \in T$ and each $\lambda_1, \dots, \lambda_p \in \mathbb{R}$*;

(iii) $\{b_n\}$ *and* $\{a_n n^{-1}(\log\log n)^{-1/2}\}$ *are nonincreasing sequences*;

(iv) $$\lim_{\gamma\to 1+} \limsup_{n\to\infty} \sup_{m:\, n\le m\le \gamma n} a_n^{-1}|a_m - a_n| = 0$$

and $\lim_{\gamma\to 1+} \limsup_{n\to\infty} \sup_{m:\, n\le m\le \gamma n} b_n^{-1}|b_m - b_n| = 0$;

(v) *for each* $t \in T$, $\lim_{u\to 1-} \rho(ut, t) = 0$, *where*

$$\rho^2(s,t) := R(s,s) + R(t,t) - 2R(s,t).$$

(vi) $$\lim_{\tau\to 0+} \limsup_{n\to\infty} \sup_{t\in T} n^{-1}a_n^2 E\Big[\big(g(X, b_n t) - E[g(X, b_n t)]\big)^2 \times I_{G(X, b_n)\ge \tau(\log\log n)^{-1/2} a_n^{-1} n}\Big) = 0,$$

where $G(x) := \sup_{t\in T}|g(x, b_n t) - E[g(X, b_n t)]|$;

(vii) *there are positive constants* r_1 *and* r_2, *such that*

$$\sum_{j=2}^{\infty} \sup_{r_1(\log j)^{-1}\le r\le r_2} e^{-r_1 r^{-1}2^j} \Pr\Big\{G(X, b_{2^j}) \ge r2^j(\log j)^{1/2} a_{2^j}^{-1}\Big\} < \infty;$$

(viii) (T, ρ) *is totally bounded*;

(ix) $a_n(2\log\log n)^{-1/2}\sup_{t\in T}|(P_n - P)g(\cdot, b_n t)| \to_{\mathrm{P}} 0$;

(x) $\limsup_{\delta\to 0}\limsup_{n\to\infty}\sup_{\rho(s,t)\le\delta} a_n^2 n^{-1}\operatorname{Var}(g(X, b_n s) - g(X, b_n t)) = 0$.

Then, with probability 1,

$$\big\{(2\log\log n)^{-1/2} a_n(P_n - P)g(\cdot, b_n t) : t\in T\big\}, \qquad n \ge 1, \tag{3.1}$$

is relatively compact in $l_\infty(T)$ *and its limit set is contained in the unit ball of the r.k.h.s. of the covariance function* $R(\cdot,\cdot)$.

PROOF. It follows from Lemma 5, using the method in the proof of Theorem 3.1 in Arcones (1994b). So, we omit the proof. □

Since the first element of the Hoeffding decomposition of $I_{h\le t}$ can be difficult to find, we first give a sharp upper bound.

THEOREM 7. *Let* $h: S^m \to \mathbb{R}$ *be a symmetric measurable function and let* $0 < p < 1$. *Suppose that there is a real number* ξ_0 *such that* $H(\xi_0) = p$, H *is continuous in a neighborhood of* ξ_0, H *is differentiable at* ξ_0, $H'(\xi_0) > 0$ *and there exists a finite constant* b *such that*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + O\big((t-\xi_0)^2\big) \tag{3.2}$$

as $t \to \xi_0$. *Then*

$$\limsup_{n\to\infty} (n/2\log\log n)^{3/4}\big|H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)\big| \le l \quad a.s.,$$

where $l = 2^{1/2}3^{-3/4}m^{3/2}(\operatorname{Var}(g(X, \xi_0)))^{1/4}$.

PROOF. Take

$$M > m\big(H'(\xi_0)\big)^{-1}\big(\operatorname{Var}\big(g(X, \xi_0)\big)\big)^{1/2}.$$

Let $b_n = (2 \log\log n/n)^{1/2}$ and let

$$Z_n(t) := (n/2\log\log n)^{3/4}(H_n(\xi_0 + tb_n) - H(\xi_0 + tb_n) - H_n(\xi_0) + H(\xi_0)).$$

By Theorem 4.7 in Arcones and Giné (1995),

$$\sup_{|t|\le M} (n/2\log\log n)^{3/4}\left|U_n\pi_{k,m}\left(I_{h\le \xi_0+tb_n} - I_{h\le \xi_0}\right)\right| \to 0 \quad \text{a.s.} \tag{3.3}$$

We claim that by Theorem 6, with probability 1,

$$\left\{(n/2\log\log n)^{3/4} m(P_n - P)(g(\cdot,\xi_0 + b_n t) - g(\cdot,\xi_0)) \colon |t| \le M\right\} \tag{3.4}$$

is relatively compact and its limit set is contained in the unit ball of the r.k.h.s. of the mean-zero Gaussian process $\{Z(t)\colon |t| \le M\}$ having covariance $E[Z(t)Z(s)] = 2^{-1}m^2H'(\xi_0)(|s| + |t| - |s-t|)$ for each $s, t \in \mathbb{R}$; that is, the limit set is contained in

$$K_0 := \left\{(\gamma(t))_{|t|\le M} \colon \gamma(0) = 0 \text{ and } \int_{-M}^{M} (\gamma'(t))^2\, dt \le m^2 H'(\xi_0)\right\}. \tag{3.5}$$

We are applying Theorem 6 with $T = [-M, M]$, $R(s,t) = E[Z(t)Z(s)]$ and $a_n = n^{3/4}(2\log\log n)^{-1/4}$. We have that

$$\limsup_{n\to\infty} b_n^{-1} m^2 \operatorname{Var}\left(\sum_{j=1}^{p} \lambda_j\left(g(X,\xi_0 + b_n t_j) - g(X,\xi_0)\right)\right) \le \sum_{j,k=1}^{p} \lambda_j\lambda_k R(t_j, t_k),$$

because

$$b_n^{-1} m^2 \left(E\left[\sum_{j=1}^{p} \lambda_j\left(g(X,\xi_0 + b_n t_j) - g(X,\xi_0)\right)\right]\right)^2 \to 0$$

and

$$b_n^{-1} m^2 E\left[\left(\sum_{j=1}^{p} \lambda_j\left(g(X,\xi_0 + b_n t_j) - g(X,\xi_0)\right)\right)^2\right] \le b_n^{-1} m^2 E\left[\left(\sum_{j=1}^{p} \lambda_j\left(I_{h\le \xi_0 + b_n t_j} - I_{h\le\xi_0}\right)\right)^2\right] \to \sum_{j,k=1}^{p} \lambda_j\lambda_k R(t_j, t_k).$$

The rest of the hypotheses in Theorem 6 are either trivial or can be checked similarly to the conditions already checked. So, (3.4) and (3.5) follow. By (3.3)–(3.5),

$$\{Z_n(t)\colon |t| \le M\} \tag{3.6}$$

is a.s. relatively compact and its limit set is contained in K_0. Let

$$Z_n(\infty) := (n/2\log\log n)^{1/2}(\xi_n - \xi_0) \tag{3.7}$$

and let $Z(\infty)$ be a Gaussian r.v. with mean 0 and variance

$$m^2(H'(\xi_0))^{-2}\operatorname{Var}(g(X,\xi_0))$$

and independent of the process $\{Z(t): |t| \le M\}$, let $T_M^* = [-M, M] \cup \{\infty\}$ and let $\rho(s,t) = \|Z(t) - Z(s)\|_2$, where $s, t \in T_M^*$. Using the method in Remark 10 in Arcones (1994a) (applying Kolmogorov's exponential inequalities), it is easy to see that for any $t_1, \dots, t_p \in T_M^*$, $\{(Z_n(t_1), \dots, Z_n(t_p))\}$ is a.s. relatively compact and its limit set is contained in the unit ball of $(Z(t_1), \dots, Z(t_p))$. Since $\{Z_n(t): |t| \le M\}$ satisfies a type of compact law of the iterated logarithm,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{\substack{|t_1|,|t_2|\le M\\ \rho(t_1,t_2)\le\delta}}|Z_n(t_1) - Z_n(t_2)| = 0 \quad \text{a.s.}$$

and $([-M, M], \rho)$ is totally bounded. So, it is also true that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\sup_{\substack{t_1,t_2\in T_M^*\\ \rho(t_1,t_2)\le\delta}}|Z_n(t_1) - Z_n(t_2)| = 0 \quad \text{a.s.}$$

and (T_m^*, ρ) is totally bounded. So, by Lemma 5 the process $\{Z_n(t): t \in T_M^*\}$ is a.s. relatively compact and its limit set is contained in the reproducing kernel Hilbert space of $\{Z(t): t \in T_M^*\}$; that is, it is contained in

$$K_1 := \Big\{(\gamma(t))_{t\in T_M^*}: \gamma(0) = 0 \quad \text{and}$$
$$(H'(\xi_0))^{-1}\int_{-M}^{M}(\gamma'(t))^2\,dt$$
$$+(\operatorname{Var}(g(X,\xi_0)))^{-1}(H'(\xi_0))^2(\gamma(\infty))^2 \le m^2\Big\}.$$

By Theorem 4.1 in Arcones (1993),

$$(n/2\log\log n)^{1/2}(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)) \to 0 \quad \text{a.s.} \tag{3.8}$$

From this, the law of the iterated logarithm of $\{Z_n(t): t \in T_M^*\}$ and composition

$$\big\{(n/2\log\log n)^{3/4}(H_n(\xi_n) - H(\xi_n) + H_n(\xi_0) - H(\xi_0))\big\}$$

is a.s. relatively compact and its limit set is contained in

$$\Big\{\gamma(v): \gamma(0) = 0 \quad \text{and} \quad H'(\xi_0)^{-1}\int_{-M}^{M}(\gamma'(t))^2\,dt$$
$$+(\operatorname{Var}(g(X,\xi_0)))^{-1}(H'(\xi_0))^2 v^2 \le m^2\Big\}. \tag{3.9}$$

By the argument in Proposition 1.1 in Deheuvels and Mason (1992), the set in (3.9) is $[-l, l]$, where $l = 2^{1/2}3^{-3/4}m^{3/2}(\operatorname{Var}(g(X,\xi_0)))^{1/4}$. □

Next, we see how the order $(n/2\log\log n)^{3/4}$ can be attained in the Bahadur–Kiefer representation of U-quantiles.

THEOREM 8. *Let* $h: S^m \to \mathbb{R}$ *be a symmetric measurable function. Let* $0 < p < 1$. *Suppose that*:

(i) *there is a real number* ξ_0 *such that* $H(\xi_0) = p$, H *is continuous in a neighborhood of* ξ_0, H *is differentiable at* ξ_0, $H'(\xi_0) > 0$ *and there exists a finite constant* b *such that*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + b(t - \xi_0)^2 + o\big((t - \xi_0)^2\big)$$

as $t \to \xi_0$;

(ii) *there is a real number* β *such that*

$$\lim_{\varepsilon \to 0+} \varepsilon^{-1} E\Big[|g(X, \xi_0 + \varepsilon t) - g(X, \xi_0 + \varepsilon s)|^2\Big] = \beta^2 |t - s|$$

for each $t, s \in \mathbb{R}$.

(iii) $$\lim_{\epsilon \to 0+} \epsilon^{-1/2} E[(g(X, \xi_0 + \epsilon t) - g(X, \xi_0))g(X, \xi_0)] = 0$$

for each $t \in \mathbb{R}$.

Then, with probability 1,

$$\Big\{(n/2 \log\log n)^{3/4}\big(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)\big)\Big\}$$

is relatively compact and its limit set is $[-l, l]$, *where*

$$l = 2^{1/2} 3^{-3/4} m^{3/2} \beta \big(H'(\xi_0)\big)^{-1/2} \big(\mathrm{Var}(g(X, \xi_0))\big)^{1/4}.$$

In particular,

$$\limsup_{n \to \infty} (n/2 \log\log n)^{3/4}\big(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)\big) = l \quad a.s.$$

PROOF. Take $M > m(H'(\xi_0))^{-1}(\mathrm{Var}(g(X, \xi_0)))^{1/2}$. The proof is similar to the proof of Theorem 7. We have that by Theorem 3.1 in Arcones (1994b)

$$\Big\{(n/2 \log\log n)^{3/4} m(P_n - P)\Big(g\big(\cdot, \xi_0 + t(2 \log\log n/n)^{1/2}\big) - g(\cdot, \xi_0)\Big) : |t| \le M\Big\}$$

is a.s. relatively compact and its limit set is

$$\left\{(\gamma(t))_{|t| \le M} : \gamma(0) = 0 \quad \text{and} \quad \int_{-M}^{M} (\gamma'(t))^2 \, dt \le m^2 \beta^2\right\}.$$

So, the arguments in Theorem 7 imply the result. □

Theorem 8 applies to $h(x_1, \dots, x_m) = \max_{1 \le i \le m} x_i$, if there exists ξ_0 with $(F(\xi_0))^m = p$, where $F(t)$ is the distribution function of X_i, F is second differentiable at ξ_0 and $F'(\xi_0) > 0$. In this case

$$l = 2^{1/2} 3^{-3/4} m p^{(4m-3)/4m} \big(1 - p^{1/m}\big)^{1/4}.$$

THEOREM 9. *Let* $h\colon S^m \to \mathbb{R}$ *be a symmetric measurable function. Let* $0 < p < 1$. *Suppose that*:

(i) *there is a real number* ξ_0 *such that* $H(\xi_0) = p$, H *is continuous in a neighborhood of* ξ_0, H *is differentiable at* ξ_0, $H'(\xi_0) > 0$ *and there exists a finite constant* b *such that*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + b(t - \xi_0)^2 + o\big((t - \xi_0)^2\big)$$

as $t \to \xi_0$;

(ii) *there is a real number* β *such that*

$$\lim_{\varepsilon \to 0+} \varepsilon^{-2} \operatorname{Var}(g(X, \xi_0 + \varepsilon t) - g(X, \xi_0 + \varepsilon s)) = \beta^2 (t - s)^2$$

for each $t, s \in \mathbb{R}$;

(iii) $\lim_{\varepsilon \to 0+} \varepsilon^{-2} E\Big[|g(X, \xi_0 + \varepsilon M)$

$$-g(X, \xi_0 - \varepsilon M)|^2 I_{|g(X, \xi_0 + \varepsilon M) - g(X, \xi_0 - \varepsilon M)| \ge \tau}\Big] = 0$$

for each $M, \tau > 0$;

(iv) *there is a real number* α *such that*

$$\lim_{\epsilon \to 0+} \epsilon^{-1} \operatorname{Cov}(g(X, \xi_0 + \epsilon t) - g(X, \xi_0), g(X, \xi_0)) = \alpha t$$

for each $t \in \mathbb{R}$;

(v) *there is a* $\delta > 0$ *such that for any* $|t - \xi_0| < \delta$ *and for any combinations* $i_1 < \cdots < i_m$ *and* $j_1 < \cdots < j_m$ *such that* $\{i_1, \dots, i_m\} \ne \{j_1, \dots, j_m\}$, *we have that*

$$\Pr\big\{h(X_{i_1}, \dots, X_{i_m}) = h(X_{j_1}, \dots, X_{j_m}) = t\big\} = 0.$$

Then, with probability 1,

$$\big\{(n/2 \log\log n)(H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0))\big\}, n \ge 1,$$

is relatively compact and its limit set is $\{\lambda_1 x_1^2 + \lambda_2 x_2^2\colon x_1^2 + x_2^2 \le 1\}$, *where* λ_1 *and* λ_2 *are in* (2.15) *and* (2.16).

PROOF. Take $M > m(H'(\xi_0))^{-1}(\operatorname{Var}(g(X, \xi_0)))^{1/2}$. Let

$$Z_n(t) := (n/2 \log\log n)\Big(H_n\big(\xi_0 + t(2\log\log n/n)^{1/2}\big) - H\big(\xi_0 + t(2\log\log n/n)^{1/2}\big) - H_n(\xi_0) + H(\xi_0)\Big)$$

for $|t| \le M$ and let $Z_n(\infty)$ as in (3.7). By the method in Theorem 8, we get that $\{Z_n(t)\colon t \in T_M^*\}$ is a.s. relatively compact and its limit set is

$$\big\{(\gamma(t))_{t \in T_M^*}\colon \gamma(t) = t u_1 \text{ for } |t| \le M \quad \text{and} \quad \gamma(\infty) = u_2, \text{ where } (u_1, u_2) \in K\big\},$$

where K is the unit ball of the reproducing kernel Hilbert space of the random vector (Y_1, Y_2), where Y_1 and Y_2 are jointly normal random variables with mean zero and

$$E\big[Y_1^2\big] = c_{1,1}, E\big[Y_2^2\big] = c_{2,2} \quad \text{and} \quad E[Y_1 Y_2] = c_{1,2},$$

where $c_{1,1}$, $c_{2,2}$ and $c_{1,2}$ are as in (2.17). By composition

$$\Big\{(n/2\log\log n)\big(H_n(\xi_n) - H(\xi_n) - H_n(\xi_0) + H(\xi_0) + b(\xi_n - \xi_0)^2\big)\Big\},\quad n \geq 1,$$

is a.s. relatively compact and its limit set is

$$\{u_1u_2 + bu_2^2 \colon (u_1, u_2) \in K\}.$$

Therefore,

$$\{(n/2\log\log n)(H_0'(\xi_0)(\xi_n - \xi_0) + H(\xi_n) - H(\xi_0))\}, n \geq 1,$$

is a.s. relatively compact and its limit set is

$$L := \{-u_1u_2 - bu_2^2 \colon (u_1, u_2) \in K\}.$$

Now,

$$K = \{(E[Y_1(b_1Y_1 + b_2Y_2)], E[Y_2(b_1Y_1 + b_2Y_2)]) \colon \|b_1Y_1 + b_2Y_2\|_2 \leq 1\}.$$

If $c_{1,1}c_{2,2} - c_{1,2}^2 = 0$, then $Y_1 = c_{1,1}^{1/2}g$ and $Y_2 = \text{sign}(c_{1,2})c_{2,2}^{1/2}g$, where g is a standard normal random variable. Thus,

$$K = \{(E[Y_1xg], E[Y_2xg]) = \{(c_{1,1}^{1/2}x, \text{sign}(c_{1,2})c_{2,2}^{1/2}x) \colon x^2 \leq 1\}$$

and

$$L = \{-(c_{1,2} + bc_{2,2})x^2 \colon x^2 \leq 1\}.$$

We have that

$$\begin{aligned}\{\lambda_1x_1^2 &+ \lambda_2x_2^2 \colon x_1^2 + x_2^2 \leq 1\}\\ &= \{(c_{1,2} + bc_{2,2})^- x_1^2 - (c_{1,2} + bc_{2,2})^+ x_2^2 \colon x_1^2 + x_2^2 \leq 1\} = L,\end{aligned}$$

where λ_1 and λ_2 are as in (2.15) and (2.16). So, the claim follows in this case. Assume that $c_{1,1}c_{2,2} - c_{1,2}^2 > 0$. Letting $Z_1 = c_{2,2}^{-1/2}(c_{1,1}c_{2,2} - c_{1,2}^2)^{-1/2}(c_{2,2} \times Y_1 - c_{1,2}Y_2)$ and $Z_2 = c_{2,2}^{-1/2}Y_2$, we have that Z_1 and Z_2 are independent standard normal random variables and

$$\begin{aligned}K &= \{(E[Y_1(d_1Z_1 + d_2Z_2)], E[Y_2(d_1Z_1 + d_2Z_2)]) \colon \|d_1Z_1 + d_2Z_2\|_2 \leq 1\}.\\ &= \Big\{\Big(c_{2,2}^{-1/2}\big(c_{1,1}c_{2,2} - c_{1,2}^2\big)^{1/2}d_1 + c_{2,2}^{-1/2}c_{1,2}d_2c_{2,2}^{1/2}d_2\Big) \colon d_1^2 + d_2^2 \leq 1\Big\}.\end{aligned}$$

From this and the fact that

$$\begin{aligned}-\Big(c_{2,2}^{-1/2}\big(c_{1,1}c_{2,2} - c_{1,2}^2\big)^{1/2}d_1 + c_{2,2}^{-1/2}c_{1,2}d_2\Big)c_{2,2}^{1/2}d_2 - b\big(c_{2,2}^{1/2}d_2\big)^2\\ = -a_1d_1d_2 - a_2d_2^2,\end{aligned}$$

where $a_1 = (c_{1,1}c_{2,2} - c_{1,2}^2)^{1/2}$ and $a_2 = c_{1,2} + bc_{2,2}$, we get that

$$L = \{-a_1d_1d_2 - a_2d_2^2 \colon d_1^2 + d_2^2 \leq 1\}.$$

If $x_1 = q_1^{-1}a_1d_1 + q_1^{-1}(a_2 + (a_1^2 + a_2^2)^{1/2})d_2$ and $x_2 = q_2^{-1}a_1d_1 + q_2^{-1}(a_2 - (a_1^2 + a_2^2)^{1/2})d_2$, where q_1 and q_2 are as in (2.18), then

$$-a_1d_1d_2 - a_2d_2^2 = \lambda_1x_1^2 + \lambda_2x_2^2$$

and

$$d_1^2 + d_2^2 = x_1^2 + x_2^2.$$

Therefore,

$$L = \left\{\lambda_1 x_1^2 + \lambda_2 x_2^2 \colon x_1^2 + x_2^2 \le 1\right\}$$

and the result follows. □

Theorem 9 applies to kernels like $h(x_1, x_2) = x_1 + x_2$, $h(x_1, x_2) = |x_1 - x_2|^r$, $r > 0$, and $h((x_1, y_1), (x_2, y_2)) = (y_2 - y_1)/(x_2 - x_1)$, under some regularity conditions. We omit the details.

Using arguments similar to those used above, it is easy to obtain the following theorems.

THEOREM 10. *Let $\{X_j\}_{j=1}^{\infty}$ be a sequence of i.i.d. r.v.'s with values in a measurable space $(S, \mathscr{S})$. Let (T, ρ) be a pseudometric space. Let $g\colon S \times T \to \mathbb{R}$ be a function such that $g(\cdot, t)\colon S \to \mathbb{R}$ is a measurable function for each $t \in T$. Let $\{b_n\}$ be a sequence of real numbers from the interval $(0, 1]$ and let $\{a_n\}$ be a sequence of positive real numbers. Suppose that*:

(i) *there is a scalar product defined for each $t \in T$ and each $0 \le u \le 1$, so that $ut \in T$*;

(ii) $\limsup_{n\to\infty} \sup_{t\in T} a_n^2 n^{-1} \operatorname{Var}(g(X, b_n t)) < \infty$;

(iii) $\{b_n\}$ *and* $\{a_n n^{-1} (\log\log n)^{-1/2}\}$ *are nonincreasing sequences*;

(iv) $\limsup_{n\to\infty} \sup_{n \le m \le 2n} a_n^{-1} a_m < \infty$;

(v) *there are positive constants r_1 and r_2, such that*

$$\sum_{j=2}^{\infty} \sup_{r_1(\log j)^{-1} \le r \le r_2} e^{-r_1 r^{-1}} 2^j \Pr\left\{G(x, b_{2^j}) \ge r 2^j (\log j)^{1/2} a_{2_j}^{-1}\right\} < \infty,$$

where $G(x) := \sup_{t\in T} |g(x, b_n t) - E[g(X, b_n t)]|$;

(vi) $(2 \log\log n)^{-1/2} a_n \sup_{t\in T} |(P_n - P) g(\cdot, b_n t)| = O_{\mathrm{P}}(1)$.

Then there is a finite constant c such that

$$\limsup_{n\to\infty} \sup_{|t|\le M} (2\log\log n)^{-1/2} a_n \left|(P_n - P) g(\cdot, b_n t)\right| \le c \quad a.s.$$

THEOREM 11. *Let $h\colon S^m \to \mathbb{R}$ be a symmetric measurable function. Let $0 < p < 1$. Suppose that*:

(i) *there is a real number ξ_0 such that $H(\xi_0) = p$, H is continuous in a neighborhood of ξ_0, H is differentiable at ξ_0, $H'(\xi_0) > 0$ and there exists a constant b such that*

$$H(t) = H(\xi_0) + H'(\xi_0)(t - \xi_0) + b(t - \xi_0)^2 + o\left((t - \xi_0)^2\right)$$

as $t \to \xi_0$;

(ii) *there is a* $1 \le v \le 2$ *such that*

$$\limsup_{\varepsilon \to 0} |\varepsilon|^{-v} E\Big[|g(X, \xi_0 + \varepsilon) - g(X, \xi_0)|^2 \Big] < \infty.$$

(iii) *there is a* $\delta > 0$ *such that for any* $|t - \xi_0| < \delta$ *and for any combinations* $i_1 < \cdots < i_m$ *and* $j_1 < \cdots < j_m$ *such that* $\{i_1, \dots, i_m\} \ne \{j_1, \dots, j_m\}$, *we have that*

$$\Pr\Big\{ h\big(X_{i_1}, \dots, X_{i_m}\big) = h\big(X_{j_1}, \dots, X_{j_m}\big) = t \Big\} = 0.$$

Then there is a finite constant c such that

$$\limsup_{n \to \infty} (n/2 \log\log n)^{(v+2)/4} |H_n(\xi_0) - H(\xi_0) + H'(\xi_0)(\xi_n - \xi_0)| \le c \quad a.s.$$

REFERENCES

ALEXANDER, K. S. (1987). Central limit theorems for stochastic processes under random entropy conditions. *Probab. Theory Related Fields* **75** 351–378.

ANDERSEN, N. T. and DOBRIĆ, V. (1987). The central limit theorem for stochastic processes. *Ann. Probab.* **15** 164–177.

ARCONES, M. A. (1993). The law of the iterated logarithm for *U*-processes. *J. Multivariate Anal.* **47** 139–151.

ARCONES, M. A. (1994a). Some strong limit theorems for *M*-estimators. *Stochastic Process. Appl.* **53** 241–268.

ARCONES, M. A. (1994b). On the law of the iterated logarithm for triangular arrays of empirical processes. Unpublished manuscript.

ARCONES, M. A. (1995). Some remarks on the weak convergence of stochastic processes. *Statist. Probab. Lett.* To appear.

ARCONES, M. A. and GINÉ, E. (1993). Limit theorems for *U*-processes. *Ann. Probab.* **21** 1494–1542.

ARCONES, M. A. and GINÉ, E. (1995). On the law of the iterated logarithm for canonical *U*-statistics and processes. *Stochastic Process. Appl.* **58** 217–245.

ARONSZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.

BICKEL, P. J. and LEHMANN, E. L. (1979). Descriptive statistics for nonparametric models. IV. Spread. In *Contributions to Statistics* (*Hájek Memorial Volume*) (J. Jurečková, ed.) 33–40. Academia, Prague.

CARROLL, R. J. (1978). On almost sure expansions for *M*-estimates. *Ann. Statist.* **6** 314–318.

CHOUDHURY, J. and SERFLING, R. J. (1988). Generalized order statistics, Bahadur representations and sequential nonparametric fixed-width confidence intervals. *J. Statist. Plann. Inference* **19** 269–282.

CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

DEHEUVELS, P. and MASON, D. M. (1992). A functional L.I.L. approach to pointwise Bahadur–Kiefer theorems. In *Probability in Banach Spaces* **8** 255–266. Birkhäuser, Boston.

DUDLEY, R. M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330.

DUDLEY, R. M. (1984). A course on empirical processes. *Lecture Notes in Math.* **1097** 1–142. Springer, New York.

DUDLEY, R. M. (1992). Fréchet differentiability, *p*-variation and uniform Donsker classes. *Ann. Probab.* **20** 1968–1982.

DUDLEY, R. M. (1994). The order of the remainder in derivatives of composition and inverse operators for *p*-variations norms. *Ann. Statist.* **22** 1–20.

GEERTSEMA, J. C. (1970). Sequential confidence intervals based on rank tests. *Ann. Math. Statist.* **41** 1016–1026.

GIJBELS, I., JANSSEN, P. and VERAVERBEKE, N. (1988). Weak and strong representations for trimmed U-statistics. *Probab. Theory Related Fields* **77** 179–194.

HELMERS, R., JANSSEN, P. and SERFLING, R. (1988). Glivenko–Cantelli properties of some generalized empirical DF's and strong convergence of generalized L-statistics. *Probab. Theory Related Fields* **79** 75–93.

HODGES, J. L., JR. and LEHMANN, E. (1963). Estimates of location based on ranks tests. *Ann. Math. Statist.* **34** 598–611.

HOFFMANN-JØRGENSEN, J. (1984). *Stochastic Processes on Polish Spaces*. Aarhus Univ. Mat. Inst. Various Publications Series **39**. Aarhus Univ.

JANSSEN, P., SERFLING, R. and VERAVERBEKE, N. (1984). Asymptotic normality for a general class of statistical functions and applications to measure of spread. *Ann. Statist.* **12** 1369–1379.

JUREČKOVÁ, J. (1980). Asymptotic representation of M-estimators of location. *Math. Operationsforsch Statist. Ser.* **11** 61–73.

JUREČKOVÁ, J. and SEN, P. K. (1987). A second order asymptotic distributional representation of M-estimators with discontinuous functions. *Ann. Probab.* **15** 814–823.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

KUELBS, J. (1976). A strong theorem for Banach space valued random variables. *Ann. Probab.* **4** 744–771.

LEHMANN, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, San Francisco.

MARCUS, M. B. and PISIER, G. (1981). *Random Fourier Series with Applications to Harmonic Analysis*. Princeton Univ. Press.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

SERFLING, R. J. (1984). Generalized L-, M- and R-statistics. *Ann. Statist.* **12** 76–86.

SHI, Z. (1995). On the uniform Bahadur–Kiefer representation. *Math. Methods Statist.* **4** 39–55.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712-1082
E-MAIL: arcones@math.utexas.edu